

We claim:

1. A computerized method for translating and transforming a heterogeneous program comprising:

obtaining a binary for a component in the heterogeneous program;

determining a plurality of basic blocks for the component, wherein the basic blocks comprise a code block and a data block;

translating each platform-specific instruction in the code block into an intermediate representation instruction to create a intermediate representation of the code block;

creating an intermediate representation of the data block;

determining a procedure within the component, wherein the procedure comprises the code block and data block;

creating an intermediate representation of the procedure from the intermediate representations of the code block and the data block;

annotating the intermediate representation of the procedure with symbol information for the code block; and

creating an intermediate representation of the component from the intermediate representation of the procedure.

2. The computerized method of claim 1, further comprising:

creating an intermediate representation of the heterogeneous program from the intermediate representation of the component.

3. The computerized method of claim 1, further comprising:
transforming the intermediate representation of the program based on user input.
4. The computerized method of claim 3, wherein the transformation comprises:
instrumenting the intermediate representation of the program.
5. The computerized method of claim 3, wherein the transformation comprises:
optimizing the intermediate representation of the program.
6. The computerized method of claim 1, further comprising:
using the symbol information in the intermediate representation of the procedure to
define an address space for the heterogeneous program; and
translating each intermediate representation instruction into a platform-specific
instruction in the address space.
7. The computerized method of claim 6, further comprising:
outputting the symbol information used to annotate the intermediate representation of
the procedure.
8. The computerized method of claim 6, further comprising:
outputting emitted block information about the intermediate representations of the
code block and the data block.

9. The computerized method of claim 8, wherein obtaining the binary for a component comprises obtaining the emitted block information output by a previous iteration of translating and transforming the heterogeneous program and wherein the emitted block information is employed in creating the intermediate representations of the code block and data block.

10. The computerized method of claim 6, wherein each intermediate representation instruction is translated into the platform-specific instruction based on user input.

11. The computerized method of claim 1, wherein translating each platform-specific into an intermediate representation instruction comprises:

replacing a common platform-specific instruction with a platform-neutral intermediate representation instruction; and

replicating a complex platform-specific instruction in an intermediate representation instruction.

12. The computerized method of claim 1, wherein the intermediate representation of the program is arranged in a hierarchy and the hierarchy comprises:

a code block element referencing each intermediate representation instruction in the intermediate representation of the code block;

a data block element;

a procedure element referencing the code block element and the data block element;
and

a component element referencing to the procedure element.

13. The computerized method of claim 1, wherein the code block element references a single intermediate representation instruction for multiple instances of a platform-specific instruction in the code block.

14. The computerized method of claim 13, further comprising:
associating a hash value with the single intermediate representation instruction so that the multiple instances of the platform-specific instructions hash to the hash value.

15. A computer-readable medium having computer-executable instructions to cause a computer to perform a reader method comprising:

obtaining a binary for a component;

analyzing the binary; and

building an intermediate representation for the component.

16. The computer-readable medium of claim 15, further comprising:
building an intermediate representation of a program containing the component.

17. The computer-readable medium of claim 15, wherein analyzing the binary comprises:

performing code discovery on the binary to determine a plurality of basic blocks; and
establishing the block relationships among the basic blocks.

18. The computer-readable medium of claim 15, wherein building the intermediate representation for the component comprises translating every instruction in the component into an intermediate representation for the instruction.

Sub D'7
197 A computer-readable medium having stored thereon an intermediate representation instruction data structure comprising:

an opcode field containing data representing an opcode for the intermediate representation instruction;

a destination operand field containing data representing a destination operand for the opcode identified by the opcode field;

a first source operand field containing data representing a first source operand for the opcode identified by the opcode field;

a second source operand field containing data representing a second source operand for the opcode identified by the opcode field;

a size field containing data representing a length for the instruction identified by a combination of the opcode field, the destination operand field, the first source operand field and the second source operand field; and

a mode field containing data representing an addressing mode for the instruction identified by the combination.

20. The computer-readable medium of claim 19, further comprising:
a shared field containing data representing a flag that indicates the instruction identified by the combination occurs multiple times in a binary.

21. The computer-readable medium of claim 19, further comprising:
an inserted field containing data representing a flag that indicates the instruction identified by the combination was inserted into a binary as a result of user action.

22. The computer-readable medium of claim 19, further comprising:
a destination operand type field containing data representing a type for the destination operand identified by the destination operand field;
a first source operand type field containing data representing a type for the a source operand identified by the a first source operand field; and
a second source operand type field containing data representing a type for the source operand identified by the a second source operand field.

23. The computer-readable medium of claim 19, further comprising:
a signature field containing data representing a particular architecture when the instruction identified by the combination is a complex instruction.

24. The computer-readable medium of claim 23, further comprising:

at least one architecture specific field containing data representing additional information contained in the complex instruction identified by the combination.

Sub D² 25. A computer-readable medium having stored thereon a hierarchical data structure for an intermediate representation of a heterogeneous program comprising:

a component data structure for a component in the heterogeneous program, the component data structure comprising a procedure field containing data representing a pointer to a procedure data structure for a procedure in the component;

the procedure data structure comprising a first block field containing data representing a pointer to a code block data structure for a code block in the procedure identified by the procedure field; and

an instruction data structure comprising an instruction field containing data representing a pointer to an instruction data structure for an instruction in the code block identified by the first block field.

Sub E 26. The computer-readable medium of claim 25, wherein the procedure data structure further comprises a second block field containing data representing a pointer to a data block data structure for a data block in the procedure identified by the procedure field.

27. The computer-readable medium of claim 25, wherein the procedure data structure further comprises a symbol field containing data representing symbol table information for the procedure identified by the procedure field.

28. The computer-readable medium of claim 25, wherein the code block data structure comprises a header field containing structure information for the block identified by the first block field.

29. The computer-readable medium of claim 28, wherein the structure information comprises:

a size;

an address; and

a set of information flags.

30. The computer-readable medium of claim 25, further comprising:

a program data structure for the heterogeneous program, the program data structure comprising a component field containing data representing a pointer to the component data structure.

31. A computerized system comprising:

a processing unit;

a system memory coupled to the processing unit through a system bus;

a computer-readable medium coupled to the processing unit through a system bus; and

a translation and transformation system executed from the computer-readable medium

by the processing unit, wherein translation and transformation system causes the processing

05043805-063099

Subl
a3

unit to translate a platform-specific binary into a plurality of intermediate representation instructions.

32. The computerized system of claim 31, further comprising:

an application program interface executing from the computer-readable medium by the processing unit and coupled to the translation and transformation system such that input received by the translation and transformation system from the application program interface instructs the translation and transformation system to further cause the processing unit to transform the plurality of intermediate representation instructions.

33. The computerized system of claim 32, wherein the translation and transformation system further causes the processing unit to translate the plurality of intermediate representation instructions as transformed into a modified platform-specific binary.

34. The computerized system of claim 33, wherein the translation and transformation system further causes the processing unit to translate the modified platform-specific binary into a modified plurality of intermediate representation instructions for further transformation.

35. A computer-readable medium having computer-executable instructions stored thereon for performing a method comprising:

translating a platform-specific binary into a plurality of intermediate representation instructions.

36. The computer-readable medium of claim 35, further comprising:
transforming the plurality of intermediate representation instructions in accordance
with input parameters.

37. The computer-readable medium of claim 36, further comprising:
translating the plurality of intermediate representation instructions as transformed into
a modified platform-specific binary.

38. The computer-readable medium of claim 37, further comprising:
translating the modified platform-specific binary a modified plurality of intermediate
representation instructions for further transformation.

39. The computer-readable medium of claim 35, further comprising:
translating the plurality of intermediate representation instructions into a new version
of the platform-specific binary.

40. The computer-readable medium of claim 39, further comprising:
translating the new version of the platform-specific binary into a new version of the
plurality of intermediate representation instructions.

41. A computer-readable medium having computer-executable instructions stored thereon for performing a method comprising:

iterating an intermediate representation of a heterogeneous program through a computerized system to create a plurality of new versions of the heterogeneous program.

42. The computer-readable medium of claim 41, further comprising:

manipulating the intermediate representation using data input into the computerized system to create the plurality of new versions of the heterogeneous program.

43. The computer-readable medium of claim 41, further comprising:

terminating the iterating of the intermediate representation based on data input into the computerized system.